- She applied an ANOVA the model for which included treatment, site and the treatment-by-site interaction. The assumption of normality for this model were however violated based on the Shapiro-Wilk test (p=0.017) and she therefore carried out a log transformation of the data. On re-applying the Shapiro-Wilk test to the log-transformed data the assumptions of normality was no longer violated (p=0.7365). An ANOVA on the log-transformed data revealed a p-value for the overall comparison of 0.0034. The subsequent comparison of each GHB group with placebo revealed the following

Comparison	p-value
GHB 3 g vs placebo	0.6358
GHB 6 g vs placebo	0.0772
GHB 9 g vs placebo	0.0021

- She had earlier also performed a Wilcoxon Rank Sum test. The p-value for the overall comparison of the 4 treatment groups, using the latter non-parametric test was 0.0101. She then compared each GHB group with placebo and the p-values for each of these comparisons was as follows

Comparison	p-value
GHB 3 g vs placebo	0.4684
GHB 6 g vs placebo	0.1450
GHB 9 g vs placebo	0.0033

- Thus, according to the protocol-specified primary efficacy analysis, and the sponsor's analysis, only the 9 g/day dose showed a statistically significant superiority to placebo in reducing the total number of cataplexy attacks.
- The evidence for efficacy at the 6 g/day dose appeared marginal and analysis-dependent.
- There was no definite evidence that GHB was efficacious in treating complete cataplexy attacks, the most serious form of cataplexy. However the mean (and median) frequency of such attacks in both treatment groups was small as was the absolute change in frequency from baseline to endpoint; a trend to a treatment effect may have been seen.

6.14.2 Secondary Efficacy Measures

- In this application the sponsor has sought a claim for Xyrem® in treating daytime sleepiness accompanying narcolepsy.
- The secondary efficacy measures used to assess daytime sleepiness included the Epworth Sleepiness Scale, the frequency of sleep attacks (inadvertent naps) during the day and the duration of daytime sleep attacks (inadvertent naps)
- On the sponsor's analysis, a nominally statistically significant superiority ($p < 0.05$) of GHB over placebo was seen on the Epworth Sleepiness Scale, the frequency of daytime sleep attacks and the duration of daytime sleep attacks, as measures of excessive daytime sleepiness. However given that there were 12 secondary efficacy measures, only the analysis of the Epworth Sleepiness Scale was still statistically significant after adjustment for multiple comparisons.

- On the sponsor's analysis, the pairwise comparisons for the Epworth Sleepiness Scale indicated that only the 9 g/day dose of GHB was superior to placebo
- Dr Sharon Yan, Agency Statistical Reviewer finds the analysis of secondary efficacy measures for this study problematical for the following reasons
  - There are many secondary efficacy measures
  - The methods of analysis were not stated in detail a priori
  - In specific reference to excessive daytime sleepiness, as measured by the Epworth Scale
    - She applied the protocol-specified ANOVA model to the original scale
    - After the treatment-by-site interaction was found not to be significant, it was removed from the model after which the residuals were no longer normally distributed, even after log transformation
    - She therefore performed a Kruskall-Wallis test. The overall p-value obtained for the GHB-placebo comparison was then 0.0109. As noted earlier, the Epworth Sleepiness Scale was one of 10 secondary efficacy measures and the overall p-value for this measure did not achieve statistical significance when a Bonferroni adjustment was made.
- The results of this study, based on the sponsor's analysis, nevertheless, do provide at least some support for the efficacy of GHB in a dose of 9 g/day in treating excessive daytime sleepiness in narcolepsy.

6.14.3 Influence Of Stimulant Drugs On Efficacy

6.14.3.1 Background

At the request of the Biopharmaceutics staff at the Agency, the following request was passed on to the sponsor on April 4, 2001

"The clinical study database should be investigated further to ascertain potential pharmacodynamic interactions in narcoleptic patients with other commonly used drugs in this patient populations."

The structure of the sponsor-proposed analysis of these interactions was discussed between the Division and sponsor at a teleconference on 4/18/01. The sponsor suggested the following, which was acceptable to the Division:

- The analysis would focus on the differences in observed effects, as they related to both safety and efficacy in narcoleptic patients

  and would compare the following groups

- Patients who received sodium oxybate alone

  Patients who received a selected concomitant medication alone

  Patients who received a combination of sodium oxybate and a selected concomitant medication

The above analysis is the basis for an additional submission dated May 4, 2001 which is summarized here as well as in my NDA Safety Review.

6.14.3.2 Sponsor's Methods

6.14.3.2.1 General Observations

- Narcoleptic patients commonly use the following classes of medications to treat that disorder
  - Stimulants (e.g., methylphenidate, dextroamphetamine, methamphetamine, pemoline, modafinil) to treat excessive daytime sleepiness
  - Tricyclic antidepressants and selective serotonin re-uptake inhibitors to treat REM dissociation phenomena: cataplexy, hypnagogic hallucinations and sleep paralysis
- The entire NDA database did not include a trial specifically designed to investigate the potential pharmacodynamic interactions between Xyrem® and medications commonly used in patients with narcolepsy. Nevertheless, for analysis purposes all clinical trials in the database were examined
- However, in only the OMC-GHB-2, Lammers and Scrima trials was it possible to compare the following groups in a controlled setting
  - Patients who received sodium oxybate alone
  - Patients who received a selected concomitant medication alone
  - Patients who received a combination of sodium oxybate and a selected concomitant medication

  Even in the setting of these 3 controlled trials
  - Stimulants were the only medication class on which such an analysis could be performed
  - Both the Scrima and Lammers trials were not suitable for the analysis on account of a small sample size and variable use of stimulants

6.14.3.2.2 Stimulant Use In OMC-GHB-2

Of the 136 patients enrolled in this trial

- 115/136 (84.6%) maintained stable doses of stimulants during the trial
- 21/136 (15.4%) did not take stimulants
- The distribution of these patients by treatment group is below

Treatment Group	Placebo	3 g/day	6 g/day	9 g/day	Total
Number treated with stimulants	28	31	26	30	115
Number not treated with stimulants	6	3	7	5	21
Total	34	34	33	35	136

Of those taking stimulant drugs

- 41 were taking amphetamines
- 55 were taking methylphenidate
- 25 were taking pemoline
- Some patients took more than 1 stimulant drug
- The distribution of these patients by treatment group is in the following table

	Treatment Group	Placebo	3 g/day	6 g/day	9 g/day	Total
Amphetamines	Number not treated with amphetamine	26	23	20	26	95
	Number treated with amphetamine	8	11	13	9	41

Methylphenidate	Number not treated with methylphenidate	17	19	24	21	81
	Number treated with methylphenidate	17	15	9	14	55
Pemoline	Number not treated with pemoline	29	28	26	28	111
	Number treated with pemoline	5	6	7	7	25

6.14.3.2.3 Analysis Of Effects Of Stimulant Drugs On Efficacy

- The 2 outcome variables chosen for the analysis were
  - The frequency of all cataplexy attacks
  - Daytime sleepiness as measured by the Epworth Sleepiness Scale
- Descriptive statistics were calculated for each of the outcome variables, for each stimulant and for patients not taking stimulants, by treatment group
- Analysis of the total number of cataplexy attacks (after log transformation) and the change in Epworth scores was accomplished using ANCOVA: the model included baseline value of the variable being analyzed (the covariate) and site and treatment as terms.
- Separate analyses were performed for each treatment group of patients, based on type of stimulant used, and any stimulant use
- Adjustments were made for multiple comparisons using the Dunnett-Hsu procedure
- An additional analysis was performed to assess the possible interaction between stimulant use and GHB treatment. This used the same ANCOVA model as above with 2 additional terms: stimulant use (yes or no) and the stimulant-by-treatment interaction. This analysis was performed for each of the stimulants above and for the stimulant group as a whole

6.14.3.3 Results: Effects Of Stimulants On Efficacy Of GHB

The differences among treatment groups were consistent between those patients taking stimulants and those not taking stimulants; this was determined using the additional ANCOVA model which included the stimulant-by-treatment-group interaction. For each stimulant and each efficacy variable this interaction was not statistically significant.

Full tables describing the analysis are in the submission. I have not reproduced them here but they appear to confirm the sponsor's conclusions

6.14.3.4 Results: Effects Of Stimulants On Safety Of GHB

These results are described in the NDA Safety Review

6.14.3.5 Sponsor's Conclusions

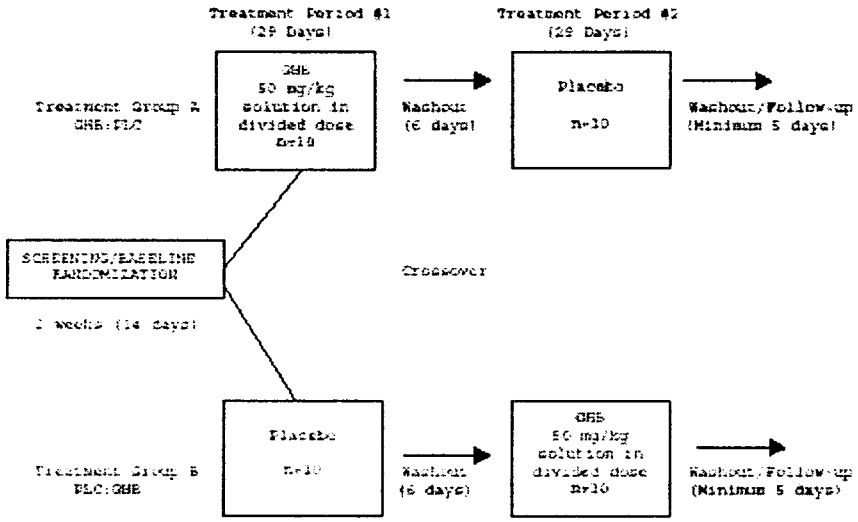
- In the analysis of cataplexy and daytime sleepiness there was no evidence of pharmacodynamic interaction between sodium oxybate treatment and concomitant stimulants
- In the analysis of adverse events there was only one body system (Digestive) in which a very weak signal of difference between those treated with GHB and methylphenidate and GHB alone was detected

- Sleep patterns identified on the polysomnogram
- Average number of REM onsets by Multiple Sleep Latency Test

7.3 Design

Randomized, double-blind, placebo-controlled, single-center, cross-over study comparing the effect of GHB 50 mg/kg total daily dose in with placebo in 20 patients with narcolepsy.

A schematic outline of the study design is presented in the figure below which I have copied from this submission.



Randomization was to be such that half of the men and half of the women participating in the study would receive GHB during the first 29-day double-blind treatment period, and placebo during the second. The remaining patients were to receive GHB first and placebo later.

7.4 Duration

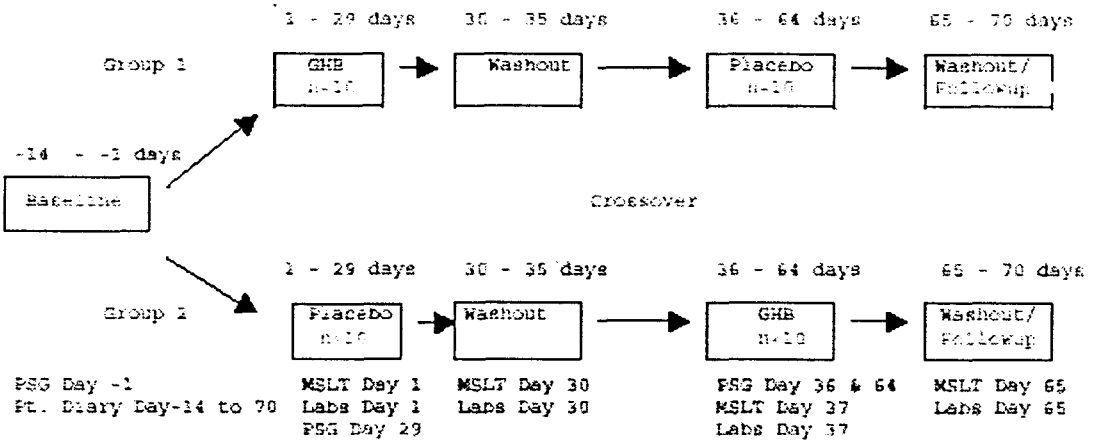
4 weeks of double-blind treatment during each cross-over period

7.5 Dosage

During each period of double-blind treatment, each participating patient was to take

GHB 25 mg/kg at bedtime, and about 3 hours later (total dose: 50 mg/kg/day)
OR
Matching placebo

Study Schedule



In the above schematic
MSLT stands for Multiple Sleep Latency Test
PSG stands for polysomnogram

- Screening assessments consisted of a medical history, sleep disorders interview, physical examination, an overnight polysomnogram and Multiple Sleep Latency Test (the last 2 tests were for diagnostic purposes only)
- Polysomnograms (overnight) were performed on the last day of the baseline period, the first night (Day 1) of each treatment period and the last night (Day 29) of each treatment period (note that this differs somewhat from what is stated in the above schematic)
- Multiple Sleep Latency Tests were performed at some time during the baseline period, on the morning and afternoon after the Day 1 overnight polysomnogram and on Day 29 of each treatment period. On each day the test was performed 5 times, each time over a 20 minute period. The timing of these epochs was as follows: 0800, 1000, 1200, 1400 and 1600 hours.
- A nighttime sleep log and daytime questionnaire were maintained by each patient throughout each study period (baseline, treatment period 1, washout, treatment period 2 and washout/follow-up). These assessed the following
  Nighttime sleep onset latency, nighttime arousals, total sleep time, feelings on awakening, number of sleep attacks, number of cataplexy attacks, number of naps, methylphenidate use, sleep/awake patterns, and mood
- Pre-sleep questionnaires were administered before each overnight polysomnogram. These assessed the following:
  General physical and mental health, food intake, activities on test days, sleep quality, and feelings
- Post-sleep questionnaires were administered on the morning after each overnight polysomnogram, after completion of the Multiple Sleep Latency. Test. They assessed the following
  Sleep quality, and feelings

- Adverse events were to be recorded continually based on patient interviews, diaries and telephone contacts. The period of observation for these began at the time of obtaining informed consent and extended through at least the first 5 days of the washout/follow-up period at the conclusion of Treatment Period 2
- Standard safety laboratory tests (hematology, chemistry and thyroid functions) were checked at the beginning and end of each treatment period
- During each polysomnogram the following additional items were monitored:
  - Cardiac rhythm by electrocardiogram
  - Respiration using either intercostal EMG or diaphragm bellows
  - Ventilation by nasal O2 and/or thermocouples
  - Blood arterial O2 saturation by ear oximetry
  - Cataplexy by EEG, chin-EMG, EOG and electrocardiogram
- Women of child-bearing potential were to have a urine pregnancy test performed 30-45 days and 7 days prior to the baseline polysomnogram.

7.11 Efficacy Outcome Measures

The efficacy outcome measures below are those listed in the study report. The primary outcome measures match those in the final analysis plan submitted by Dr Scrima in December1986, clarified further in March 1987. The secondary efficacy measures only partly match those listed in the final analysis plan.

7.11.1 Primary Efficacy Measures

- Number of cataplexy attacks per day
- Objective daytime sleepiness as measured by the Multiple Sleep Latency Test* Sleepiness Index

7.11.2 Secondary Efficacy Measures Based On Patient Diaries

- The number of awakenings from sleep at night
- The number of sleep attacks during the day
- Methylphenidate use (mg/day)
- Patient sense of alertness on awakening on a scale from 1 (feeling active, alert, vital and wide awake) to 8 (asleep): this is also referred to as the Stanford Sleepiness Scale

  Feeling active, vital, alert, or wide awake: 1
  Functioning at high levels, but not at peak; able to concentrate: 2
  Awake, but relaxed; responsive but not fully alert: 3
  Somewhat foggy, let down: 4
  Foggy; losing interest in remaining awake; slowed down: 5
  Sleepy, woozy, fighting sleep; prefer to lie down: 6
  No longer fighting sleep, sleep onset soon; having dream-like thoughts: 7
  Asleep: 8
- Mood in the morning recorded on a scale from –10 (extremely negative) to +10 (extremely positive) with 0 representing a neutral mood
- Mood in the evening recorded on a scale from –10 (extremely negative) to +10 (extremely positive) with 0 representing a neutral mood

7.11.3 Secondary Efficacy Measures Based On Polysomnograms

- Sleep efficiency (%)
- Sleep latency (minutes)

- Stage I (%)
- Stage 2 (%)
- Stage 3 (%)
- Stage 4 (%)
- Stage shifts (the number of times during sleep that sleep transitions from one sleep stage to another sleep stage)
- REM sleep (%)
- REM latency (minutes)
- Number of awakenings

*7.11.4 Secondary Efficacy Measures Based On Multiple Sleep Latency Test**

7.11.5 Number of REM onsets

7.11.6 Safety Outcome Measures

Adverse events
Changes in clinical laboratory parameters (adverse events and safety laboratory parameters)

*The **Multiple Sleep Latency Test** consisted of 5 separate sessions on a single day, each session lasting 20 minutes. Measurements in minutes derived from the test included the following:

- Latency to sleep onset (time from lights out to the first epoch of sleep)
- Latency to REM sleep (time from the beginning of sleep onset to the first epoch of REM sleep)
- Average latency to sleep onset (total latencies to sleep/number of naps)
- Number of naps with REM across five naps
- Sleepiness Index = 100 – (5 x average latency to sleep)

7.12 Analysis Plan

Defining the analysis plan actually used for the study took considerable effort

7.12.1 General Considerations

7.12.1.1 Efficacy Analyses

- All randomized patients were to be included in the primary and secondary efficacy analyses, except that patients missing an entire treatment period were not included in the post-treatment analysis.
- Patients included in the analysis of the effects of GHB/placebo withdrawal during the washout and washout/follow-up periods must have had both baseline data and at least 3 of the 5 days following treatment periods for each period
- Missing data were substituted for 1 or 2 weeks during the treatment periods or 1 or 2 days during the washout or washout/follow-up periods. Missing data substitutions were made only for the diary data
- All efficacy analyses were to be based on change of data from baseline to post-treatment for each study period

7.12.1.2 Safety Analysis

Any patient who received study medication was to be included in the safety analysis

7.12.1.3 Study Periods

These were defined as follows

- The baseline period was the 2-week pre-treatment period over which patient diaries were recorded. The 5-days preceding the first dose of study medication were considered an adequate baseline for data collected on patients' daily subjective assessments
- GHB/placebo were administered in Treatment Periods 1 and 2 (see schematic in Section 7.3)
- The washout period was the 5-day period after Treatment Period 1
- The follow-up period (also referred to as the washout/follow-up period) was the 5-day period after Treatment Period 2

7.12.1.4 Others

- Sleep stages were converted to percentages of sleep time for analysis purposes
- Diary dates were converted to treatment day as follows
  - The date in the diary for the last day of the baseline period was the first treatment day
  - The date in the diary for the last day of treatment for each treatment period was the day after treatment was stopped
- The criterion for a statistically significant difference was $p < 0.05$ (2-sided?)

7.12.2 Demographic And Baseline Characteristics

- Age, weight, age at diagnosis, and number of sleep and cataplexy attacks were analyzed using a 2-factor ANOVA. The effects in this model were sequence group, gender and the interaction of gender and sequence group
- The distributions of patients with or without histories of hypnagogic hallucinations or sleep paralysis was tested for independence from gender and sequence group using contingency table methods
- Only patients with baseline data who were included in the post-treatment analysis were analyzed for baseline comparability (of narcolepsy-related parameters).
- Note that the above methods for analyzing baseline and demographic characteristics were not specified prospectively.

7.12.3 Primary Efficacy Parameters

- The primary efficacy parameters consisted of the change from baseline scores for the following: the number of daily cataplexy occurrences and the sleepiness index measured from the multiple sleep latency test.
- Change from baseline for the frequency of cataplexy attacks was to be measured as per the final prospectively-designated Scrima analysis plan submitted 12/4/86 as follows
  - The mean daily number of cataplexy attacks was to be calculated from patient diaries for the following periods: the last 2 weeks of the baseline period, the last 2 weeks of the GHB treatment period and the last 2 weeks of the placebo treatment period

- The change in mean daily cataplexy frequency from the last 2 weeks of the baseline period to the last 2 weeks of the GHB treatment period and to the last 2 weeks of the placebo treatment period was then calculated

- The sponsor states that in Dr Scrima's final analysis plan, the "data called for analyzing only Weeks 1 and 4." What is stated in Dr Scrima's submission of March 6, 1987, is that 2 within patient factors would be included in the ANOVA model: substance and time (1st week versus 4th week for the patient measures, and 1st day versus 28th day for the sleep study measures). If the intention was to use the mean change from baseline to Weeks 1 and 4 in cataplexy frequency as a primary efficacy measure, it is not clearly stated which of these 2 mean changes (i.e., baseline to Week 1 or baseline to Week 4) would be considered the primary efficacy parameter. In the first of 2 publications describing this study (Scrima et al 1989, see Section7.1) the sponsor states that the change in mean daily cataplexy frequency from baseline to Week 1 and baseline to Week 4 were both used.
- At any rate, the post-hoc [redacted] analysis (see Section 7.12.7) is what was described in the study report. In the study report the following appear to have been done in analyzing the change in frequency of cataplexy attacks
  - The baseline mean daily cataplexy frequency was calculated based on the last 5 days of the baseline period (this is not specified in any prospective version of the analysis plan; all such versions called for the baseline frequency to be calculated using the last 14 days of the baseline period)
  - The mean daily cataplexy frequency was calculated for each of the 4 weeks of treatment with GHB or placebo.
  - The overall mean daily frequency of cataplexy attacks was calculated for each treatment period in their entirety.
  - The change from baseline for each treatment modality was then calculated based on the difference between the mean daily number of cataplexy attacks during the baseline period and the mean daily number of cataplexy attacks overall during the GHB and placebo treatment periods. The 2 treatments were then compared
- The change from baseline for the Multiple Sleep Latency Test Sleepiness Index was to be calculated, as per the final prospectively-designated Scrima analysis plan submitted 12/4/86, based on the difference in scores from the Multiple Sleep Latency Test performed on the last day of the baseline period to the last day of the GHB treatment period, and from the last day of the baseline period to the last day of the placebo treatment period. However the analysis presented in the study report appears to imply that the primary analysis of the Sleepiness Index was based upon the difference in mean scores between baseline and the entire period of treatment with either GHB or placebo.
- A multifactorial repeated measures ANOVA was used for this analysis; the "original" model submitted in the final prospective Scrima analysis plan included one between-subject factor (treatment order) and 2 within-subject factors [treatment and time (week)]. In addition to the above factors the following interactions were to be tested: treatment x order; treatment x time; and treatment x substance x time (null hypotheses are stated). The [redacted]

—— analysis (see Section 7.12.7) added a between-patient factor of gender, and all possible interactions with gender; this analysis also pooled various interaction terms related to the within-patient error term.

- Since the term week was frequently significant either as part of a main effect or interaction, further repeated-measures analyses for the individual weeks were performed to support the overall analyses
- In the final prospective Scrima analysis plan, the effect of GHB withdrawal was to be evaluated using a paired t-test contrasting the washout week following withdrawal of GHB with the washout week following withdrawal of placebo. In the actual analysis the effect of GHB withdrawal was assessed as follows
  - The analysis was confined to diary-derived data since only diary-based data were recorded during the washout and washout/follow-up periods
  - One analysis compared the 2 treatment sequences over the initial washout period, in regard to change from baseline. The analysis was based on a repeated-measures ANOVA. The model had a single within-patient factor, days, and two between-patient factors, sequence group and gender. Separate univariate supportive analyses were performed for washout days 1 to 5 with sequence group, gender and their interaction as factors. The intercept was tested in each model to identify departure from baseline
  - A second analysis compared the washout from treatment in Period 1 and the washout/follow-up period after Treatment Period 2. A repeated-measures ANOVA was performed on the change from baseline data. There were two between-patient factors, treatment group and gender, and two within-patient factors, Day 1 to 5 of washout or follow-up

7.12.4 Secondary Efficacy Parameters

- These were analyzed using methods largely similar to the primary efficacy analysis
- Conventional levels of statistical significance (i.e., $p < 0.05$) were retained for the comparisons (Dr Scrima, had in the final prospective analysis plan, stated that the a Bonferroni correction would not be used but the results would have to be interpreted with caution unless replicated)

7.12.5 Safety Parameters

- The GHB and placebo treatment periods were compared in regard to the incidence of adverse events and changes in clinical laboratory results
- Treatment-related adverse events were grouped based on body system, severity and relationship to study drug.
- The clinical laboratory parameters were examined with respect to reported values that were outside the normal range.

7.12.6 Sample Size Rationale

A formal a priori sample size calculation was not performed "as this was the first clinical trial examining GHB in the treatment of narcolepsy"

7.12.7 [redacted] Analysis

This analysis, performed, as noted earlier, in 1992, differed from the original analysis plan as follows

- An analysis of data for Weeks 1, 2, 3, and 4 was performed for the frequency of cataplexy attacks, instead of what was specified prospectively.
- The original analysis plan addressed only the analysis of male patients (it stated that since the male narcoleptic group was reaching completion at that time, the first reports were to based on a single gender group and gender was not to be included as a factor in the analysis). At the time of the [redacted] analysis, data from both the male and female patients were available and the model used at that time to analyze data from this study included a between-patient factor of gender, as well as all possible interactions with gender
- The model used in the original analysis plan included various interaction terms related to the within-patient error term. The [redacted] analysis pooled these interaction terms so as to simplify the model.

7.13 Protocol Amendments

They were no formal protocol amendments and no comprehensive final version of the protocol. Changes to the protocol have been inferred from correspondence between Dr Scrima and the Agency.

7.14 Efficacy Results

7.14.1 Patient Disposition

- 20 patients (10 men and 10 women) were enrolled in and completed the study
- Of these 10 patients (5 men and 5 women) were randomized to the GHB:placebo sequence and 10 patients (5 men and 5 women) were randomized to the placebo:GHB sequence
- All 20 patients completed the study

7.14.2 Protocol Deviations

- In the case of one 16-year old patient randomized to the placebo:GHB sequence, neither the patient nor her parents signed the original informed consent document; her parents did sign the informed consent attachment for minors and/or fertile females
- 3 patients (1 in the GHB:placebo sequence and the remaining 2 in the placebo:GHB sequence) had a baseline sleepiness index that fell below the minimum required for study entry
- 1 patient randomized to the GHB:placebo sequence took a diuretic throughout the study
- 1 patient randomized to the GHB:placebo sequence continued taking propranolol 40 mg daily for hypertension throughout the study. Propranolol has apparently been used off-label for treating narcolepsy in doses of 80-480 mg/day

- 2 patients (#s 6 and 13; one in each treatment sequence) had no diary record for cataplexy attacks during the pre-treatment period; in one of these instances (a patient, # 13, in the placebo:GHB sequence) the technician noted that the patient had had frequent cataplexy attacks and recorded that she had had 10 cataplexy attacks during the 2-week pre-treatment period, the minimum required to qualify for the study. Another patient in the GHB:placebo sequence (# 11) had attacks that were too frequent to count during the baseline period, and was therefore recorded by the technician to have 10 attacks so as to enable her to enter the study
- 1 patient in the placebo:GHB sequence got drunk during the second treatment period
- Another patient in the placebo:GHB sequence had an extended baseline period because he was initially taking > 100 mg of methylphenidate daily and was slowly tapered to a dose of 30 mg daily. He also had a 4-week washout period between treatment periods. In Weeks 1 and 2 of Treatment Period 1 his methylphenidate dose exceeded protocol guidelines
- Another patient in the placebo:GHB sequence had a 4-week washout period between treatments
- 5 patients in the GHB:placebo sequence and 4 patients in the placebo:GHB sequence took methylphenidate after 6 PM
- 2 patients in the placebo:GHB sequence and 1 patient in the GHB:placebo sequence received their numbers out of sequence

7.14.3 Baseline And Other Demographic Characteristics

These are summarized in the following 2 tables which compares the 2 treatment sequences. There were 5 men and 5 women in each treatment sequence

Variable	GHB:placebo sequence Mean (SD)	Placebo:GHB sequence Mean (SD)
Age (years)	53.2 (9.0)	42.8 (14.9)
Weight (lbs)	199.8 (22.9)	165.9 (27.2)
Age at diagnosis (years)	28.1 (7.8)	25.2 (14.7)
Number of cataplexy attacks per day	3.5 (2.9)	3.0 (2.6)
Sleep attacks per day	2.6 (0.8)	3.1 (1.5)

Variable	GHB:placebo sequence Number of patients	Placebo:GHB sequence Number of patients
History of hypnagogic hallucinations	6	4
History of sleep paralysis	7	6

As the tables above indicate

- The GHB:placebo sequence had an older mean age than the placebo:GHB sequence, although the sponsor points out that the difference was not statistically significant (p = 0.089)
- The placebo:GHB sequence had a mean lower weight than the GHB:placebo sequence, and the sponsor acknowledges that the difference was statistically significant (p = 0.008)

Analyses of concomitant medications and illnesses are not provided

7.14.4 Medication Compliance

A single patient (#6 participating in the GHB:placebo sequence) failed to bring his supply of study drug to the site on the last day of Treatment Period 1. The pharmacy at the study site prepared the last dose for that study period.

No data are available for the exact dose in grams that each patient received.

7.14.5 Primary Efficacy Analysis

7.14.5.1 Cataplexy Attacks

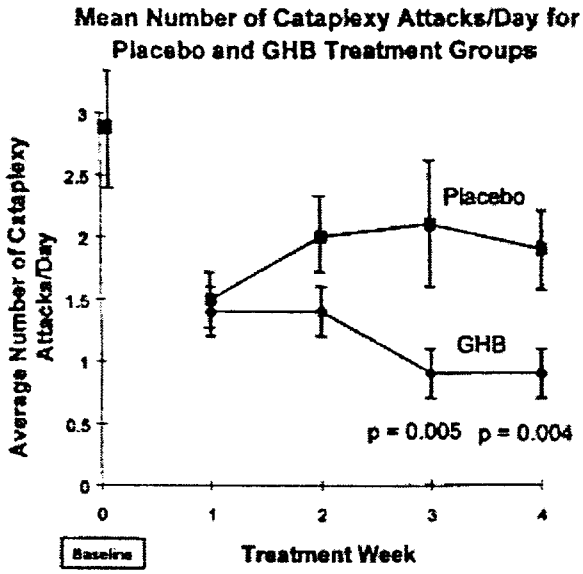
7.14.5.1.1 Baseline Comparability

An analysis of baseline comparability was initially performed, excluding patients 6 and 13 who had no diary records for the baseline period. An analysis of baseline comparability on the remaining 18 patients showed no significant differences between sequence groups, genders nor their interaction for the mean number of daily cataplexy attacks

7.14.5.1.2 Active Treatment Period

	Mean Number of Cataplexy Attacks Per Day						
	Pre-Treatment	Treatment Phase					
Treatment Group	Baseline (SE)	Week 1 (SE)	Week 2 (SE)	Week 3 (SE)	Week 4 (SE)	Overall (SE)	Baseline to Endpoint
GHB	2.6 (0.5)	1.4 (0.2)	1.4 (0.2)	0.9 (0.2)	0.9 (0.2)	1.2 (0.2)	2.9 to 1.2 (p=0.007)
Placebo		1.5 (0.2)	2.0 (0.3)	2.1 (0.4)	1.8 (0.3)	1.9 (0.3)	2.8 to 1.9 (p=0.117)
p-value between treatments	---	n.s.	n.s.	0.005	0.004	0.013	---

n.s. = not significant



The analysis of the frequency of cataplexy attacks (the primary efficacy analysis) is summarized in the above table and figure.

The p-value for the overall GHB-placebo comparison was 0.013. The same comparisons for individual weeks were statistically significant at Weeks 3 and 4, although the overall interaction of treatment with week was not statistically significant (p = 0.071). Among other comparisons made by the sponsor: 21% of GHB-treated patients and 5% of placebo-treated patients were cataplexy-free by Week 4.

No other significant main effects or interactions were identified, in particular sequence group (p = 0.775) or treatment x sequence group interaction (p = 0.713); there was thus no evidence of a carry-over effect

As the above table and figure indicate the mean decrease from baseline in the daily frequency of cataplexy attacks in each treatment group was

- 1.7 during GHB treatment (p = 0.007 for change from baseline)
- 1.0 during placebo treatment (p = 0.117 for change from baseline)

7.14.5.1.3 Washout And Washout/Follow-Up Period

Patient #s 6 and 13 were excluded from the analysis of the washout and washout/ follow-up periods, on account of a lack of baseline data and a lack of diary entries, altogether.

For the initial washout period

- Patients taking GHB had an average of 1.0 cataplexy attacks per day versus 1.7 days for those taking placebo
- During the washout period there was a mean decrease from baseline for the GHB group of 1.6 attacks per day (p = 0.054) and for placebo of 1.4 attacks per day (p = 0.065) .
- The difference between treatment groups for mean change from baseline was not statistically significant
- There was a significant increase in attack frequency for both treatment groups from Day 1 to Day 5 of the washout period. By Day 5, 25% of GHB patients and 50% of placebo patients reported as many cataplexy attacks as at baseline.

No statistically significant differences were detected between the 5 days of washout and the 5 days of washout/follow-up overall nor for the interaction with sequence group

7.14.5.2 Sleepiness Index

7.14.5.2.1 Baseline Comparability

No significant differences were detected for baseline comparability between sequence groups, genders or their interaction.

7.14.5.2.2 Active Treatment Period

The results of the analysis are summarized in the following table

Treatment Group	GHB N = 20	Placebo N = 20
Mean Baseline Sleepiness Index	88.5	
Mean Day 1 Sleepiness Index	88.6	90.9
Mean Change From Baseline At Day 1	0.1	2.4
Mean Day 29 Sleepiness Index	85.8	89.6
Mean Change From Baseline At Day 29	-2.7	1.1
Mean Overall Sleepiness Index During Treatment	87.2	90.3
Mean Overall Change From Baseline During Treatment	-1.3	1.8
GHB-Placebo Difference For Overall Treatment Effect	-3.1	
P-value for overall GHB-placebo difference	0.085	

As the table above indicates the difference between the GHB and placebo groups was not statistically significant overall. The decrease in sleepiness index with GHB did appear to be greater at Day 29 than at Day 1. There were however no statistically significant day effects ($p = 0.057$) or interaction for treatment group x gender ($p = 0.081$).

7.14.6 Analysis Of Secondary Efficacy Measures

Note that there were a total of 16 secondary efficacy measures recorded in the study report. These were in 2 categories: those based on patient diaries and those based on polysomnogram data

Note that the published reports of the Scrima Study

7.14.6.1 Secondary Efficacy Measures Based On Patient Diaries

These analyses are summarized in the next table copied from the submission. The treatment effect (GHB or placebo) listed in the table is the overall effect for 4 weeks of treatment.

The table below indicates that GHB was superior to placebo at a nominal level of statistical significance (($p = 0.042$) only in regard to subjective awakenings at night. When adjusted for multiple comparisons this difference was no longer statistically significant.

Summary of Secondary Efficacy Variables from Patient Diaries					
Variable	Baseline (mean)	Treatment Group	Overall (4 Weeks of TX)	P Value Vs. Baseline	P Value Btwn TXs
Sleep Attacks/Day	2.8	PLC	2.1	p=0.037*	0.530
		GHB	1.9	p=0.002*	
Subjective Awakenings at Night	3.0	PLC	3.7	p=0.095	0.042**
		GHB	2.4	p=0.091	
Methylphenidate (mg/day)	22.0	PLC	14.4	p=0.091	0.792
		GHB	11.4	p=0.400	
How Patients Felt Upon Awakening	2.9	PLC	2.7	p=0.472	0.219
		GHB	2.5	p=0.072	
Mood in the Morning	2.3	PLC	2.2	p=0.867	0.142
		GHB	2.8	p=0.348	
Mood in the Evening	0.9	PLC	1.5	p=0.251	0.210
		GHB	2.2	p=0.007*	

TX - Treatment
* Significant compared to Baseline
** Significant difference between treatments

7.14.6.2 Secondary Efficacy Measures Based On Polysomnograms

These analyses are summarized in the next table copied from the submission. While the GHB-placebo difference for the overall treatment effect was nominally statistically significant ($p < 0.05$) for a number of polysomnographic variables, only the effect on Stage 3 sleep (increased with GHB relative to placebo) remained so after adjustment for multiple comparisons

Summary of Secondary Efficacy Variables - PSG Results								
Variable	Baseline	TX Group	Overall Days 1 & 29	Overall Days 1 and 29 Btwn TXs	Day 1	Day 1 Btwn TXs	Day 29	Day 29 Btwn TXs
Sleep Efficiency (%)	83.5	PLC	88.1	p=0.021**	87.3	p=0.019**	88.9	p=0.120
		GHB	84.4		83.5		85.1	
Sleep Latency (min)	9.4	PLC	2.5	p=0.028**	2.5	p=0.210	2.6	p=0.164
		GHB	3.5		3.6		3.4	
Stage 1 Sleep (%)	19.9	PLC	17.0	p=0.042**	25.4	p=0.122	28.4	p=0.026**
		GHB	23.0		21.7		24.3	
Stage 2 Sleep (%)	48.7	PLC	45.6	p=0.712	46.3	p=0.881	44.4	p=0.642
		GHB	44.8		44.7		42.5	
Stage 3 Sleep (%)	3.5	PLC	2.8	p=0.001**	3.2	p=0.166	2.3	p=0.001**
		GHB	5.5		4.3		5.7	
Stage 4 Sleep (%)	4.6	PLC	4.2	p=0.321	3.7	p=0.200	4.7	p=0.691
		GHB	5.4		5.6		5.2	
Stage Shifts	112.8	PLC	122.2	p=0.306**	125.1	p=0.005**	133.2	p=0.051
		GHB	109.9		102.9		116.9	
REM Sleep (%)	19.0	PLC	20.9	p=0.425	21.9	p=0.646	19.9	p=0.537
		GHB	21.5		22.3		20.7	
REM Latency (min)	51.2	PLC	50.4	p=0.113	36.8	p=0.267	44.3	p=0.117
		GHB	34.4		37.4		35.1	
Number of Objective Awakenings	27.2	PLC	26.9	p=0.012**	25.1	p=0.042**	25.9	p=0.042**
		GHB	21.8		20.7		22.8	
		GHB	27.2		4.8		3.4	

TX - Treatment
** Significant difference between treatments

7.15 Safety Results

These are described in the NDA Safety Review

7.16 Sponsor's Conclusions

The sponsor's key conclusions may be summarized as follows:

- After 4 weeks of treatment GHB showed a statistically significant superiority to placebo in reducing the frequency of all cataplexy attacks. This effect appeared to be greatest during Weeks 3 and 4 of treatment.
- GHB did not demonstrate a statistically significant superiority to placebo on objective daytime sleepiness as measured by the Multiple Sleep Latency Test Sleepiness Index
- The sponsor has made a number of conclusions regarding efficacy based on
  - Nominally statistically significant results for the GHB-placebo comparisons for secondary efficacy measures
  - Within group changes from baseline for both primary and secondary efficacy measures

7.17 Reviewer's Comments

- The Scrima study has many deficiencies. The main ones are as follows
  - The protocol and its subsequent changes are not presented in an entirely coherent or easily understood form
  - The last prospectively-described modification of the protocol does not clearly define the endpoints for the primary efficacy analysis
  - The primary efficacy analysis described in the study report is post-hoc, and not identical to the prospectively-defined analysis plan
  - Some carry-over effects might have occurred during the initial 5-day washout period between treatment sequences (see Section 7.14.5.1.3)
- Despite the above deficiencies it does appear that, regardless of how the study endpoint was measured, GHB showed an at least nominally statistically significant superiority to placebo in reducing the total frequency of cataplexy attacks
- Dr Sharon Yan, Agency Statistical Reviewer, has conveyed the following to me
  - This study had 2 primary efficacy measures, as already noted
  - It was not specified a priori that a "win" on both measures at a $p < 0.05$ level of significance would be required to declare the study positive
  - Therefore after adjusting for multiple comparisons the level of significance for declaring a "win" on either primary efficacy measure would have to be < 0.025
  - Whereas the sponsor's analysis of cataplexy attacks showed a drug-placebo difference at a p-value of 0.013, her own analysis has yielded a p-value for the same comparison of 0.0372
  - She is therefore of the view that the study does not have adequate evidence of the efficacy of GHB in treating cataplexy
- The sponsor's analysis has revealed that GHB did not have a statistically significant superiority to placebo on the Sleepiness Index, calculated from the

Multiple Sleep Latency Test, as a measure of excessive daytime sleepiness. Dr Sharon Yan agrees with this conclusion

- Individual absolute doses (i.e., g/day) of GHB are not stated in the protocol. The total nightly dose of GHB was 50 mg/kg as per the protocol. However based on individual weight data listings in the study report
  - Weights of patients participating in the study ranged at entry from 119 to 249 lbs (54.1 to 113.2 kg)
  - Doses of GHB can therefore be estimated to have ranged from 2.7 to 5.7 g/day
  - The mean weight for patients participating in the study was 199.8 lbs (90.8 kg)
  - The mean dose of GHB might therefore be estimated as 4.5 g/day
- Note that in the 2 published reports of this study
  - There was no distinction between the primary and secondary efficacy variables
  - Each report deals with a different set of outcome variables
  - Neither report mentions the Multiple Sleep Latency Test Sleepiness Index as an outcome variable
- **Also note the Clinical Inspection Summary from this center's Division of Scientific Investigations (DSI) (see Section 15.3). The Scrima study site was inspected at the request of this Division. DSI concluded that the data from the Scrima study site were unacceptable since the drug accountability records were largely missing. It was recommended by DSI that the data from the Scrima study not be used in support of this application.**

8. Lammers Study

This study was done outside IND purview in the Netherlands. The results of this study are available both as a study report, as presented in submission # 37 under IND # —— and as an article in the medical literature (Lammers GJ et al. Gammahydroxybutyrate and Narcolepsy: A Double-Blind, Placebo-Controlled Study. Sleep 1993; 16:216-220). The analysis and conclusions presented in the study report differ somewhat from that in the publication

8.1 Objectives

To demonstrate that GHB in narcoleptics

- Has an effect on REM dissociation phenomena, and on excessive daytime sleepiness
- Has an effect on alertness during the day
- Has a mood-improving effect

8.2 Design

Randomized, double-blind, placebo-controlled, cross-over study

The study consisted of the following phases:
A **baseline** observation period lasting 1 week followed by
2 **treatment** periods each lasting 4 weeks, separated by a
A **washout** period lasting 4 weeks

Note that the last week of the washout period was considered a baseline observation period for the second treatment phase

During the initial treatment period a patient was to be randomly assigned to either GHB or placebo; during the second treatment period that patient would cross over to whatever treatment had not been administered during the first phase. The order in which the patient received either GHB or placebo was therefore randomly decided.

8.3 Inclusion Criteria

- Any race or gender
- Written informed consent
- Combination of sleep attacks during the day, and at least one of the REM dissociation phenomena (cataplexy, hypnagogic hallucinations and sleep paralysis); in case of doubt regarding the diagnosis of narcolepsy, a positive Multiple Sleep Latency Test, as recorded with a 24 hour electroencephalogram, was required

8.4 Exclusion Criteria

- Atypical narcolepsy; narcolepsy was to be considered atypical if
  - The above combination of 2 clinical criteria was absent
  - The Multiple Sleep Latency Test was negative
- Serious liver, kidney or cardiac disorders
- Pregnant patients, or those liable to become pregnant during the study
- One or more serious suicide attempts in the past

8.5 Concomitant Medication

All concomitant medications were to be continued but no alteration of dose was permitted during the study.

8.6 Dosage

GHB 30 mg/kg taken twice each night, with the initial dose at bedtime and the second dose 4 hours later

or matching placebo

The mean daily dose of GHB actually used in the study was 4.75 g/day

8.7 Schedule

- Physician assessments appear to have been planned at recruitment, at the end of the baseline period, at the end of the first treatment period, at the end of the washout period and at the end of the second treatment period (the assessment at the end of the washout period was to serve as the baseline for the second treatment period)
- Telephone contact was to be made between the study physician and patient every week
- A daily diary was to be maintained by the patient during the baseline period, during the last week of the washout period and throughout each treatment

period recording the number of cataplexy attacks, the number of daytime sleep attacks, number of awakenings at night, whether refreshed on awaking in the morning, severity of daytime sleepiness, and the number of episodes of hypnagogic hallucinations, sleep paralysis and automatic behavior

- The severity of narcolepsy was assessed at the baseline measurement using a narcolepsy questionnaire, a sample of which was supplied with this submission. A sleep score (Rechtschaffen and Kales) and a vigilance score (Simon and Schulz) were also completed at that time
- A mood rating scale was completed at baseline and during every other week of each treatment period
- A 24 hour cassette electroencephalogram recording and a Multiple Sleep Latency Test were performed at baseline, during the one-week period before the second treatment phase and at the end of each treatment phase
- The opinions of the patient (Global Therapeutic Impression) and clinician (Global Clinical Impression) as to whether improvement had occurred or not were to be obtained at the end of each treatment period.
- Safety monitoring procedures are not specified

8.8 Efficacy Measures

Note that the distinction between primary and secondary efficacy variables is not completely clear in the original protocol. The efficacy measures are listed in the original protocol in the same sequence as they are listed below, but not under separate "Primary" and "Secondary" headings. The protocol however states: "If patients show an improvement on *all* (the emphasis is mine) of the first 3 criteria, this is regarded as a positive result" and "the other variables are of secondary importance and will only be analyzed if effect is found in the primary variables". Based on these statements in the protocol, I have listed certain outcome measures under the "Primary" heading and the remainder under the "Secondary" heading

8.8.1 Primary

- Number of cataplexy attacks per week
- Global Therapeutic Impression (patient): this was scored at the end of the entire study on a hand-written sheet and at the end of each treatment period in the daily diary.
- Global Clinical Impression (clinician)

8.8.2 Secondary

- Number of sleep attacks during the day
- Feeling of sleepiness during the day (using a visual analogue scale)
- Multiple Sleep Latency Test improvement of the two shortest latencies with a minimum of 4 minutes in total
- Stability of alertness during the day (based on electroencephalogram)
- Duration of slow wave sleep
- Decrease in number of phase shifts at night
- Change in mood (?) - this is not consistently specified to be an efficacy parameter in the protocol.

The sponsor states that the secondary efficacy variables were to be analyzed only if an "effect" was found in the primary efficacy variables

8.8.3 Safety Measures

Not specified, except that adverse events were to be recorded in the daily diary supplied to participating subjects. There is no evidence from either the study report or publication that vital signs, safety laboratory tests, electrocardiograms, or physical examinations were checked and recorded.

8.9 Analysis Plan

- **The analysis plan is not presented in a well-organized manner in the version of the protocol that is in the study report, which appears to be largely, but not entirely, a description of the analysis as actually performed; the original protocol (presented in an appendix) contains only the following statements under the heading "Statistical Analysis": "Differences between placebo and gammahydroxybutyrate will be tested by means of the Wilcoxon signed rank test ($\alpha = 0.05$; two-sided). The diary data will be subjected to a trend analysis (including a rank order test for changes within the group and between groups)"**
- The median/mean score for the variables was calculated for the baseline observational period that preceded each treatment period, and for each of the weeks during the treatment periods. Next, the change from baseline to study endpoint (Week 4) during the corresponding treatment period was calculated
- The analysis of most of the primary and secondary outcome measures, including the total number of cataplexy attacks, consisted of comparing the median change from baseline to endpoint for the GHB and placebo treatment periods using the Wilcoxon Signed Rank Test with a 2-sided type 1 error of 0.05. For the total number of cataplexy attacks and daytime sleep attacks, data were analyzed for treatment and period effect using the method of Pocock
- The change in mean total and individual item scores on the mood rating scale from baseline was compared for the GHB and placebo periods using the paired t-test
- For each of the efficacy variables, an ANCOVA was also performed using a model appropriate for a cross-over design which included the following factors: treatment order, patient position within treatment order, treatment group, period (?) and baseline value for the efficacy variable; the significance of the covariate was also examined. If the usual assumptions required for ANCOVA were not satisfied, a log transformation of the data was to be considered and the residuals were to be analyzed using the Shapiro-Wilk test. If the data were not normally distributed, a non-parametric test was performed. The ANCOVA was considered an "additional" analysis and was not mentioned at all in the original protocol

- The rating scales for, and methods of analyzing, the Global Clinical Impression and Global Therapeutic Impression are not specified in the original protocol
- "Intent-to-treat" was the only dataset analyzed as per the study report: however the criteria for including patients in this dataset are not defined. In the study report it is stated that an intention-to-treat to treat analysis was also used in the publication. No specific dataset for analysis is specified in the original protocol. The number of patients in the intention-to-treat datasets used for the study report is different from that in the publication, and the further clarified by me under "Dataset Analyzed", a subheading under "Efficacy Results" below

8.10 Protocol Changes

- Daytime alertness was dropped as a secondary outcome measure, reportedly because the investigators had difficulty defining and measuring alertness during the day
- The Global Clinical Impression was dropped as a primary outcome measure as it was not apparently recorded ("the investigator only reproduced the opinion of the patient and could not translate this information into his own opinion")
- "Polysomnogram" (presumably the author is referring to 24 hour cassette electroencephalogram recording, as true polysomnographic data was not obtained, as per the study report) data were "no longer available" and were therefore not analyzed.
- **From this reviewer's perspective all elements of the analysis plan, such as ANCOVA, that were outlined in the study report, but not included in the original protocol as presented in an appendix, may be considered changes in the analysis plan**
- **This reviewer requested the sponsor to clarify whether any formal protocol amendments exist for this study. In submission # 39 under IND # ——— , dated November 6, 1998, the sponsor responded to this request and clarified that to the best of their knowledge and that of the ——— clinical contractor ——————— that assisted in the preparation of the study report for this submission, there were no protocol amendments. The sponsor also clarified that Dr Lammers decided to drop several procedures that were specified in the original protocol; I have already outlined these procedures above.**
- **In submission # 39 under IND # ——— (dated 11/6/98) the sponsor further indicated that an ANCOVA was used for the analysis of the data from this study "to maintain consistency with the statistical analysis of the OMC-GHB-2 clinical trial data".** In this connection I would point out that whereas the results of the Lammers study were first published in 1993 (the actual study period was from November 1987 to December 1988), the protocol for OMC-GHB-2 was finalized (with its last amendment) only in February 1997. Thus the ANCOVA-based analysis of the Lammers data was carried out several years after the data were first unblinded.

8.11 Efficacy Results

8.11.1 Number of Patients and Disposition

- 25 patients were enrolled in and completed both treatment periods. Their randomized distribution, according to treatment sequence, was as follows

First Treatment Period	Second Treatment Period	Number of Patients
GHB	Placebo	13
Placebo	GHB	12

- 1 patient failed to maintain a diary for the first treatment period and was therefore excluded from the intent-to-treat analysis (as well as from a separate analysis for publication). Thus 24 patients were included in the intention-to-treat analysis in the study report
- Another patient was later determined not to have narcolepsy but was included in the intention-to-treat analysis in the study report (a separate intention-to-treat analysis for publication purposes excluded this patient as well; thus the intention-to-treat analysis in the publication had 23 patients)

8.11.2 Protocol Deviations

These include only the two patients referred to above, under "Number of Patients and Disposition"

8.11.3 Dataset Analyzed

- Only an intention-to-treat analysis was carried out, although a definition of this dataset was not provided in the study report or publication
- As indicated above, 24 patients were included in the intention-to-treat analysis that was in the study report and 23 patients in the intention-to-treat analysis in the publication

8.11.4 Demographic Baseline Variables

These are indicated for each treatment order in the next table copied from the submission. "GHB - Placebo" refers to the sequence where GHB was administered first and placebo second. "Placebo - GHB" refers to the sequence where placebo was administered first and GHB second.

	GHB - Placebo (n=13)	Placebo - GHB (n=12)	Total (n=25)
Age (yrs)	41 [14][3]	39 [15]	40 [14]
Male Female	8 (32%) 5 (20%)	5 (20%) 7 (28%)	13 (52%) 12 (48%)
Weight (Kg) (n=24)[1]	79 [11]	79 [9]	79 [10]
Height (Cm) (n=24)[2]	175 [8]	175 [7]	175 [7]

[1] Weight was not recorded in one patient
[2] Height was not recorded in one patient
[3] Standard deviations are indicated in parentheses

As the table indicates an imbalance of gender was present between sequences with a greater proportion of men in the "GHB - Placebo" sequence and a greater proportion of women in the "Placebo - GHB" sequence. The other demographic variables appear to have been balanced between treatment sequences.

Since this was a cross-over study, the placebo and GHB groups were identical

8.11.5 Primary Efficacy Analysis

This section will describe the intention-to-treat analysis for the following outcome measures: total number of cataplexy attacks per week, and Global Therapeutic Impression.

8.11.5.1 Total Number of Cataplexy Attacks per Week

The following table presents the overall results of this analysis; the results of the study report analysis, and publication analysis are presented in separate rows.

Source	Treatment Group	Median/Mean of Total Number of Cataplexy Attacks per Week *			p-value for Change from Baseline to Endpoint (GHB vs placebo)
		Baseline	Endpoint	Baseline-Endpoint Change	
Study Report (n = 24)	Placebo GHB	5.53 3.99	3.01 1.47	-2.52 -2.52	0.002 (ANCOVA)
Publication (n = 21)	Placebo GHB	1.56 1.26	1.24 0.56	-0.32 -0.70	0.42 (Wilcoxon)

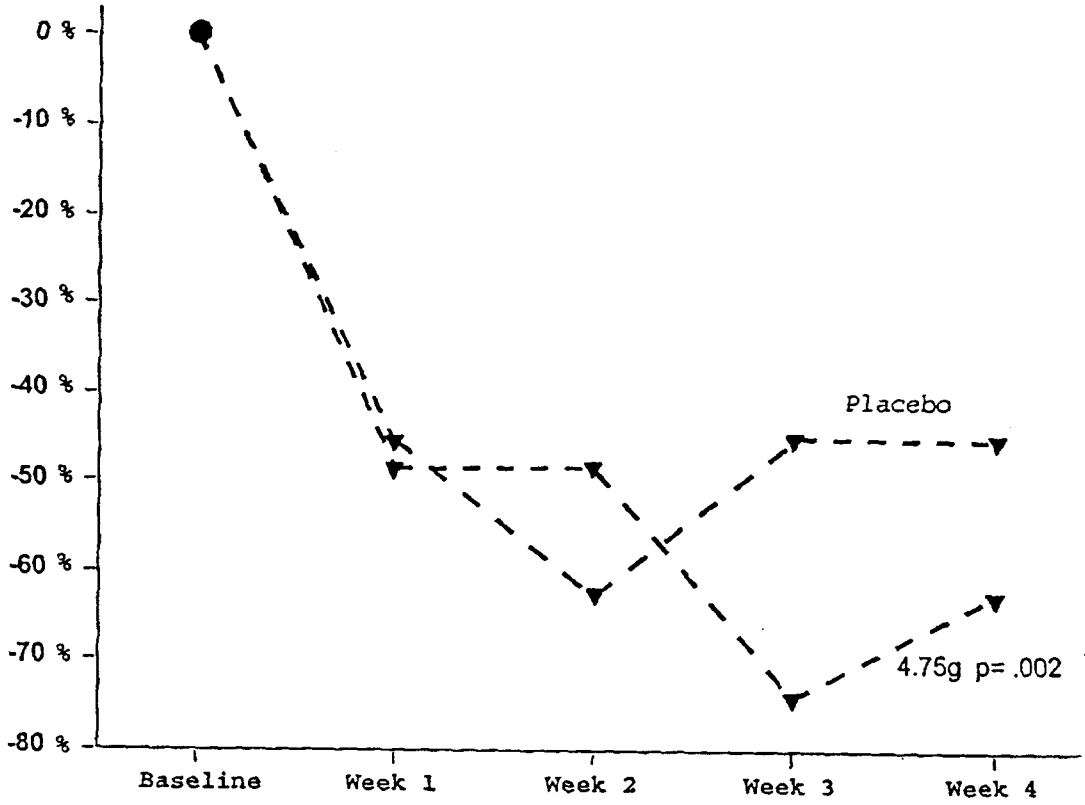
*For the study report, medians are indicated; for the publication, means are indicated

As noted earlier, the primary efficacy analysis as specified in the original protocol was to be based on the Wilcoxon signed rank test and not on ANCOVA; the latter analysis was performed after the study blind was broken.

The next table presents the full details of the ANCOVA used for this outcome measure

Distribution of residual Shapiro-Wilk test	p-values of the factors in the model		
	Treatment	Period	Covariate (Baseline value)
Normal (p = 0.88)	0.002	0.39	0.0001

The percentage change in the median total number of cataplexy attacks for each week of the study is presented graphically below, for the placebo and GHB groups.



The term "4.75 g" in the graph refers to the mean daily dose of GHB.

8.11.5.2 Global Therapeutic Impression

NOTE:

- In the publication, no distinction was made between the Global Therapeutic Impression recorded on a hand-written sheet at the end of the entire study, and that recorded in the daily diary at the end of each treatment period. However the results in the publication appear to correspond to those recorded on a handwritten sheet at the end of the study.
- As noted earlier, the original protocol does not mention how this outcome measure was to be rated or analyzed. In the publication, the Global Therapeutic Impression is described as having been rated on a 4 point scale as in the table below, although for analysis purposes the scale appears to have been dichotomized into "beneficial" or "not beneficial"; in a communication dated November 29, 1998 the sponsor has indicated that the original 4-point rating scale was collapsed so that patients reporting "not beneficial" and "possibly (presumably) beneficial" responses were considered non-responders, whereas those reporting "beneficial" and "strongly beneficial" responses were considered responders. A similar dichotomous scale has been described in the analysis detailed in the study report

No effect at all	0
Possibly beneficial	1
Beneficial	2
Strongly beneficial	3

The results of the Global Therapeutic Impression, as recorded on a handwritten sheet at the end of the entire study are presented in the table below, copied from the submission

		GHB period		
		No beneficial effect	Beneficial effect	
Placebo period	No beneficial effect	8	15	23
	Beneficial effect	1	1	2
		9	16	25

As indicated above the results are favorable for GHB in relation to placebo; this difference is statistically significant (p = 0.001; McNemar's test)

The results of the Global Therapeutic Impression, as recorded in the daily diary at the end of each treatment period are presented in the table below, again copied from the submission

		GHB period		
		No beneficial effect	Beneficial effect	
Placebo period	No beneficial effect	11	10	21
	Beneficial effect	2	2	4
		13	12	25

As indicated above the results are again favorable for GHB in relation to placebo; this difference is statistically significant (p = 0.021; McNemar's test).

In a communication dated 11/29/98 the sponsor has provided, at the request of the Agency's statisticians, evidence of why McNemar's test is an appropriate analysis in these circumstances.

APPEARS THIS WAY
ON ORIGINAL

8.11.6 Secondary Efficacy Analysis

The results for most secondary outcome measures are listed in the following table. Note that the outcome measures listed differ, in part, from those listed in the protocol. With the exception of the "Feeling Refreshed in the Morning" category a negative change represented an improvement. The "Feeling Refreshed in the Morning" measure was rated on a scale from 0 to 4 with 0 = not and 4 = very good.

Measure	Treatment Group	Median/Mean of Daily Score *			p-value for Change from Baseline to Endpoint (GHB vs placebo)
		Baseline	Endpoint	Baseline-Endpoint Change	
Severity of Daytime Sleepiness (n =24)	Placebo	1.60	1.59	-0.01	0.034
	GHB	1.60	1.28	-0.32	(Wilcoxon)
Daytime Sleep Attacks (Frequency) (n =24)	Placebo	1.83	2.14	0.31	0.0008
	GHB	2.17	1.36	-0.81	(ANCOVA)
Nocturnal Awakenings (n =24)	Placebo	2.71	3.31	0.60	0.011
	GHB	3.39	2.00	-1.39	(ANCOVA)
Refreshed in the Morning (n =24)	Placebo	1.50	1.71	0.21	0.13
	GHB	1.86	2.29	0.43	(ANCOVA)
Hypnagogic Hallucinations (n =24)	Placebo	0.14	0.00	-0.14	0.056
	GHB	0.21	0.00	-0.21	(ANCOVA)
Sleep Paralysis (n =23)	Placebo	0.03	0.01	-0.02	0.56
	GHB	0.07	0.02	-0.05	(Wilcoxon)
Automatic Behavior (n =24)	Placebo	0.27	0.30	0.03	0.23
	GHB	0.35	0.21	-0.14	(Wilcoxon)

*Median values are indicated for all items other than the severity of daytime sleepiness, sleep paralysis and automatic behavior for which means are indicated

As the above table indicates nominally statistically significant improvement with GHB relative to placebo was seen for the following measures: severity of daytime sleepiness, daytime sleep attacks and nocturnal awakenings; that for hypnagogic hallucinations approached nominal statistical significance.

According to the sponsor, "there were no differences in changes from baseline to Week 4 between the GHB period and placebo period with regard to the total score of the mood rating scale and the separate items" on that scale. The details of this analysis are not presented. All 25 patients were apparently included in the analysis of the mood rating scale.

8.12 Safety Results

These are described further in the NDA Safety Review

8.13 Investigator's Conclusions

These are summarized below

- Primary efficacy measures
  - GHB showed a statistically significant benefit relative to placebo on the Global Therapeutic Impression
  - GHB showed a statistically significant benefit relative to placebo in reducing the total number of attacks of cataplexy

- Secondary efficacy measures
  - GHB showed a statistically significant benefit relative to placebo on the following measures: severity of daytime sleepiness, number of daytime sleep attacks, and number of nighttime awakenings
  - GHB showed a "marked" benefit relative to placebo on hypnagogic hallucinations
  - A statistically significant benefit for GHB relative to placebo was not seen for the following parameters: feeling refreshed in the morning, sleep paralysis, automatic behavior, total score on mood rating scale, scores for individual items on mood rating scale

8.14 Reviewer's Comments

These comments include those that were made at the time Treatment IND # ——— was reviewed in November-December 1998, and more recent observations

- The protocol-specified analysis of one primary efficacy variable, the total number of cataplexy attacks per week, did not reveal any superiority of GHB, in the dose used, to placebo. However, this analysis, the Wilcoxon signed rank test, did not appear appropriate for a cross-over study since it did not examine cross-over and period effects
- Although not listed as a method of analysis in the protocol, an ANCOVA may have been the appropriate means of analysis for the total number of cataplexy attacks since it could adjust for the baseline differences for the GHB and placebo treatment periods. While this analysis clearly showed a statistically significant superiority of GHB over placebo, it required the assumption that the data were normally distributed. However, the data did not appear normally distributed on visual inspection, even after log transformation; although a formal test did not reject the hypothesis that the data were normally distributed, this test may not have had sufficient power to detect potentially important deviations from normality. In addition the ANCOVA assumed the absence of important carry-over effects; the data suggested that such carry-over effects might be present and that formal testing for normality might not be sufficiently powered to reject such effects. For these reasons it was unclear that the ANCOVA was an appropriate analysis.
- A standard ANOVA used for cross-over study designs was applied by Dr Sharon Yan and yielded a marginal mean difference in the number of cataplexy attacks between treatment groups of –0.359 in favor of GHB ($p = 0.1789$). Another ANOVA model yielded a p-value of 0.1233 again favoring GHB. Dr Yan considers both models to be appropriate for a study of this design
- Any analysis other than the Wilcoxon signed rank test must be considered post-hoc, although given that the protocol-specified efficacy analysis might not be appropriate, alternative analyses might.
- Dr Sharon Yan has pointed out that any conclusions as to the efficacy of GHB in treating cataplexy is analysis- and model-dependent and in the absence of a full prospectively-designated analysis plan, no conclusions can be made about the efficacy of GHB in this study.

- The other primary efficacy variable that was actually used in the analysis was the Global Therapeutic Impression which showed a statistically significant superiority for GHB over placebo. However it was unclear as to precisely what manifestations of narcolepsy this assessment was measuring
- Based on the above, the Division had expressed the opinion that it was not possible to conclude that the Lammers study was an adequate and well-controlled study contributing reliable evidence of efficacy

8.15 Reviewer's Comments About Secondary Efficacy Measures

- There were at least 9 secondary efficacy measures described in the study report. These were only partly consistent with those described in the protocol
- A nominally statistically significant ($p < 0.05$) improvement with GHB relative to placebo was seen for the following measures: severity of daytime sleepiness, frequency of daytime sleep attacks and nocturnal awakenings
- However, when adjustment was made for multiple comparisons only the effect of GHB on the frequency of daytime sleep attacks remained statistically significant ($p = 0.0008$); this was a protocol-specified outcome measure. This result was however based on ANCOVA which was not a protocol specified analysis. The severity of daytime sleepiness, measured on a visual analogue scale (which was not further described) improved more with GHB than with placebo ($p = 0.034$; Wilcoxon) but this difference did not remain statistically significant (at a $p < 0.05$ level) when adjusted for multiple comparisons
- Note that the sponsor is currently seeking a claim for the use of GHB in treating daytime sleepiness accompanying narcolepsy. The results of this study would, at best, only partly support such a claim.
- Also note that the study protocol called for continuing all concomitant medication during the study, including stimulant drugs, provided the dose was not altered. Since the study had a cross-over design, the effect of GHB on measures of daytime sleepiness is not likely to have been confounded by the concurrent use of stimulant medication unless medication changes were made during the study.

9. Study OMC-SXB-20

This was an open-label study that was intended to evaluate the effects of 4 doses of Xyrem® on sleep architecture. The study report was submitted on 12/16/00, i.e., after the original NDA submission. The sponsor desires that the results of this study be included in labeling.

A brief outline of the study protocol and a summary of efficacy data from this study are presented below.

9.1 Objectives

9.1.1 Primary

The primary objective of this study was to characterize the polysomnographic sleep architecture in narcoleptic patients at 4 GHB doses: 4.5 g, 6.0 g, 7.5 g and 9 g daily

9.1.2 Secondary

The secondary objectives of the study were to

- Assess the effect of Xyrem® on sleep as measured by the Epworth Sleepiness Scale
- Assess the effects of Xyrem® on common symptoms of narcolepsy as measured by the Narcolepsy Symptoms Assessment
- Assess EEG measures of wakefulness under soporific conditions using the Maintenance of Wakefulness Test
- Assess the safety of Xyrem®

9.2 Design/Summary of Investigational Plan

This was an open-label uncontrolled study divided into 2 phases. Stimulant medication was maintained at a constant level during the trial

9.2.1 Phase I

This phase lasted 4 weeks

- In the initial 2 weeks of this phase patients were withdrawn from tricyclic antidepressants, selective serotonin re-uptake inhibitors and hypnotics
- In the last 2 weeks of this phase patients remained free of tricyclics

An overnight polysomnogram was performed at the beginning and end of this phase. The Epworth Sleepiness Scale questionnaire was administered at about the time of each polysomnogram

9.2.2 Phase II

This phase began with the patient receiving 4.5 g of GHB nightly for the initial 4 weeks. At the end of this period the dose was increased to 6.0 g nightly and further to 7.5 g nightly and 9 g nightly at 2 week intervals. Each total nightly dose of GHB was administered in 2 equal divided doses 2.5 to 4 hours apart.

Overnight polysomnograms on the night of the first dose of Xyrem® and on the last night of each dose. The Epworth Sleepiness Scale was administered at the end of each dosing period

9.3 Duration

10 weeks

9.4 Sample Size

20-30 planned

9.5 Key Inclusion Criteria

- Informed consent
- Age ≥ 18 years
- American Sleep Disorders Association criteria for narcolepsy
- Use of stable doses of tricyclic antidepressants or selective serotonin re-uptake inhibitors for narcolepsy for at least 3 weeks. If taking stimulants must have been on a stable dose for at least 3 weeks
- If female must be

- Surgically sterile OR
- 2 years post-menopausal OR
- If of child-bearing potential must be using effective contraception and must continue this treatment during the study
- Adequate support for duration of trial

9.6 Key Exclusion Criteria

- Unstable diseases in any body system, other than narcolepsy, which would place the patient at risk or compromise the trial objectives
- Use of tricyclic antidepressants or selective serotonin re-uptake inhibitors for depression or for any indication other than narcolepsy
- History of substance abuse, as defined by DSM-IV, currently or within the past year
- History of psychiatric disorders that would preclude study participation
- Serum creatinine > 2 mg/dl; AST or ALT > 2 x upper limit of normal; serum bilirubin > 1,5 x upper limit of normal; pre-trial electrocardiogram results demonstrating a clinically significant arrhythmia or 2^{nd} or 3^{rd} degree A-V block; history of myocardial infarction within the past 6 months
- Any untreated disorder other than narcolepsy that could be considered a primary cause of excessive daytime sleepiness, including sleep apnea syndrome (criteria specified)
- Occupation requiring variable shift or routine night shift work
- Use of sodium oxybate within the preceding 30 days
- Use of any investigational drug within the preceding 30 days
- No clinically significant history of head trauma, seizure disorder or previous intracranial surgery
- Willing to not operate a car or heavy machinery if the clinical investigator feels such a restriction is warranted
- Use of medication for narcolepsy during baseline period, other than a stable dose of stimulant medication ("stable dose" defined as one without any significant change in dose for the 5 - day period just prior to the baseline period)
- Use of hypnotics, tranquilizers, antihistamines (except for the non-sedating variety of such drugs) and clonidine at the start of the baseline period.

9.7 Dosage

See Section 9.2

9.8 Outcome Measures

9.8.1 Primary Efficacy Measures

The following objective overnight polysomnogram parameters

- Wake After Sleep Onset (WASO) in minutes following the first and second dose of Xyrem and the summation
- Total Sleep Time (TST) in minutes following the first and second dose of Xyrem and the summation

- Stage 1 sleep time in minutes following the first and second dose of Xyrem and the summation
- Stage 2 sleep time in minutes following the first and second dose of Xyrem and the summation
- Stage 3 & 4 sleep time in minutes following the first and second dose of Xyrem and the summation
- Rapid Eye Movement (REM) sleep time in minutes following the first and second dose of Xyrem and the summation
- Sleep latency in minutes following the first and second dose of Xyrem
- REM sleep latency in minutes following the first and second dose of Xyrem
- Stage shifts per hour following the first and second dose of Xyrem and an average
- Total awakenings following the first and second dose of Xyrem® and the summation
- Delta power in microvolts^2/Hz following the first and second dose of Xyrem and an average

9.8.2 Secondary Efficacy Measures

- Epworth Sleepiness Scale
- Narcolepsy Symptoms Assessment
- Maintenance of Wakefulness Test

9.8.3 Safety Measures

Adverse events, safety laboratory tests, vital signs, electrocardiograms and physical examinations

9.9 Analysis Plan

- Demographic variables at baseline were summarized as follows
- Gender and race were summarized by the number of patients in each category
- Age, height and weight were summarized by descriptive statistics
- Efficacy variables were analyzed as follows
- Inferential statistics were performed for descriptive purposes only as per the sponsor
- Quantitative polysomnogram variables and the Epworth Sleepiness Scale were analyzed using 2-way ANOVA with patient and dosage as the main effects
- If a statistically significant difference was found among dose groups using ANOVA, pairwise comparisons using the least significant difference test were performed. If the assumptions for the above ANOVA were not satisfied the rank changes from baseline were analyzed using the ANOVA model. The significance of the mean change from baseline (end of Phase I) in each dose group was determined using a paired t-test or a Wilcoxon signed rank test
- For the above analysis the level of statistical significance was 0.05 (two-sided)
- Variables for the narcolepsy symptom questionnaire measured as a change from the beginning of Phase I were presented by number and percentage of patients
- Safety analyses were performed as follows
- Adverse events were summarized by body system using COSTART term and by relationship to treatment, dose and severity
- Changes from the beginning of Phase 1 to the end of the study in laboratory parameters were summarized using descriptive statistics
- Changes from the end of Phase I to the end of the study in vital signs were summarized using descriptive statistics

- Changes from the beginning of Phase I to the end of the study in electrocardiogram parameters were summarized

9.10 Results

9.10.1 Patient Disposition

- 27 patients were enrolled in the study
- 25 patients were treated with GHB
- 21 patients completed the study

9.10.2 Baseline And Demographic Characteristics

Baseline and demographic characteristics for all 25 treated patients are summarized below

Variable	Mean	Standard Deviation
Age (years)	52.6	8.77
Weight (kg)	84.2	16.36
Height (cm)	166.9	8.32

Gender: Males 28%; Females 72%
Race: Caucasian 92%; Black 8%

9.10.3 Tricyclic Antidepressants, Selective Serotonin Re-Uptake Inhibitors And Hypnotics At Baseline

These are summarized in the next table, copied from the submission.

Preferred Term	Total
Number of Patients	25 (100%)
Patients Receiving Medications	22 (88%)
Clomipramine	3 (12%)
Fluoxetine	5 (20%)
Fluvoxamine	1 (4%)
Paroxetine	2 (8%)
Protriptyline	1 (4%)
Sertraline	4 (16%)
Venlafaxine	6 (24%)

TCA = Tricyclic antidepressant. SSRI = Selective serotonin reuptake inhibitors.
All medications were completed prior to the start of treatment.

9.10.4 Protocol Deviations

These are summarized in the next table copied from the submission. The table applies to all 25 treated patients

Type of Protocol Deviation	No. of Protocol Deviations
Inclusion/exclusion criteria	6
Compliance	7
Concomitant medication	28
Study visit interval	17
Error in dosing medication	23
Efficacy measure	33
Safety measure	
Laboratory procedure	2
Other safety measure	2
Other	7
Total	125

9.10.5 Treatment Compliance

Treatment compliance at each dose level is summarized in the following table copied from the submission. Mean compliance at each dose level was high.

	Dose (g)				
Number of	4.5	6.0	7.5	9.0	Total
Patients	25	22	22	21	25
Compliance (%)					
N	25	22	22	21	25
Mean	95.9	95.5	92.7	93.3	94.9
SD	11.45	9.63	9.06	13.45	7.62
Median	100.0	95.0	95.0	97.7	95.7
Minimum					
Maximum					

9.10.6 Extent Of Exposure

The mean duration of treatment was 63.3 nights (standard deviation: 21.29)

9.10.7 Efficacy Results

The study design was open-label and uncontrolled; a design that is not accepted as a means of assessing drug efficacy. I have therefore presented only a brief summary of the efficacy results

9.10.7.1 Polysomnogram Variables

9.10.7.1.1 Slow Wave Sleep time

There were dose related increases in Stage 3 & 4 sleep time across all doses that reached statistical significance at 9.0 g per night. Stage 3 & 4 sleep changed from a mean at baseline of 3.5 minutes and increased by 0.6, 5.4, 10.7, and 23.2 minutes following the first dose of 4.5 g, 4 weeks at 4.5 g, and 2 weeks at 6.0, 7.5, and 9.0 g per night doses, respectively. The change at 9.0 g was statistically significant ($p = 0.012$, Wilcoxon signed rank test).

9.10.7.1.2 Delta power

Delta power is a derived index of all slow wave frequencies in the 0.5 to 4 Hz range for all non-REM sleep that incorporates amplitude and frequency components and has a high correlation with slow wave sleep. Delta power showed a dose related increase across all doses that was highly significant at the first dose of 4.5 g ($p < 0.001$) and after 2 weeks of dosing at 6 ($p = 0.036$), 7.5 ($p < 0.001$), and 9.0 g ($p < 0.001$) per night.

9.10.7.1.3 Nocturnal awakenings

Total nocturnal awakenings decreased in a dose related manner. Mean (SD) nocturnal awakenings were 50.2 (13.67) at baseline and decreased by 9.1 (14.2), 0.2 (14.8), 5.1 (14.8), 12.9 (13.2), and 12.4 (16.3) following the first dose at 4.5 g, 4 weeks at 4.5 g, and 2 weeks at 6.0, 7.5, and 9.0 g doses, respectively. This result was significant at the 7.5 g ($p < 0.001$) and 9 g ($p < 0.001$) dosing levels.

9.10.7.1.4 Stage 1 and 2 sleep time

There was a trend toward decrease in Stage 1 sleep and no changes in Stage 2 sleep time.

9.10.7.1.5 REM sleep time

REM sleep time increased significantly following the first 4.5 g dose and then decreased significantly in a dose related manner following 4 weeks at 4.5 g, and 2 weeks at 6.0, 7.5,and 9.0 g doses, respectively. The decrease in total REM sleep time following subsequent dosing appears to be related to the increases in slow wave sleep and delta power, a well-documented phenomenon reported in the literature.

9.10.7.1.6 REM Latency

There were no significant changes in REM latency across the doses studied.

9.10.7.1.7 Total Sleep Time

Total sleep time exhibited no dose related changes but a significant decrease after 4 weeks of 4.5 g dosing
($p = 0.041$).

9.10.7.1.8 Stage Shifts

Stage shifts per hour exhibited a trend towards decrease at doses above 4.5 g reaching significance only at the 7.5 g dose ($p = 0.002$).

9.10.7.1.9 Wake Time After Sleep Onset

Wake time after sleep onset did not change significantly from baseline and there were no discernable trends.

9.10.7.1.10 Nocturnal Sleep Latency

Sleep latency increased significantly in a dose related manner for all doses ($p < 0.05$) from 4 weeks at 4.5 g to 9.0 g. Median baseline sleep onset latency was 2 minutes and increased by 1, 1, 2, and 3.5 minutes at the 4 weeks at 4.5 g dose and the 2 weeks at 6.0, 7.5, and 9.0 g doses, respectively.

9.10.7.2 Maintenance of Wakefulness Test (MWT)

Polysomnographic measurement of daytime wakefulness utilizing the standardized MWT indicated dose related and significant increases in sleep latency. Mean (SD) baseline sleep latency was 4.5 (6.0) minutes. Mean (SD) increase in sleep latency after 4 weeks at 4.5 g sodium oxybate was 3.7 (7.7) minutes ($p=0.038$) with greater increases of 6.1 (6.8) minutes ($p < 0.001$) following two weeks after treatment to 9.0 g.

During the daytime recordings there was a dose related decrease in the percentage of patients with sleep-onset REM periods. sleep-onset REM periods were displayed in 18 of 21 patients (86%) at baseline. After 4 weeks on 4.5 g, 13 of 21 patients (62%) had sleep-onset REM periods. Six of 20 patients (30%) had sleep-onset REM periods at a dose of 9.0 g.

9.10.7.3 Epworth Sleepiness Score

There was a statistically significant dose related decrease in the Epworth Sleepiness Score for all doses ($p < 0.001$) beginning with 4 weeks at 4.5 g. The baseline median ESS was 20. The median decrease in ESS total score was 2, 3, 4, and 7 for the 4 weeks at 4.5 g, and 2 weeks at 6.0, 7.5, and 9.0 g doses, respectively. This response

was achieved despite continued concomitant treatment with stimulants at stable therapeutic doses. The decrease in ESS in this trial supports changes seen in previous sodium oxybate trials (OMC-GHB-2, OMC-GHB-3).

9.10.7.4 Narcolepsy Symptom Assessment

Patients reported dose related improvement in narcolepsy symptoms across all doses. The improvements included the number of cataplexy attacks, hypnagogic hallucinations, sleep paralysis episodes, inadvertent naps/sleep attacks during the day, awakenings at night, severity of daytime sleepiness, quality of sleep at night, ability to concentrate, and overall condition. These improvements began after 4 weeks on 4.5 g per night and continued through the end of the trial with greater improvements with increasing sodium oxybate dosage.

9.10.8 Safety Results

See summary in NDA Safety Review

9.11 Sponsor's Conclusions Regarding Efficacy

The following positive conclusions were drawn from the data

9.11.1 Polysomnogram Variables

- Xyrem treatment resulted in a dose related increase in slow wave sleep across all 4 doses reaching significance at the 9.0 g/night regimen.
- Delta power, a derived index of all slow wave signals, showed a dose related increase that was highly significant on the first night following 4.5 g as well as after 2 weeks of dosing at 6.0 g, 7.5 g, 9.0 g/night.
- A dose related decrease in the number of nocturnal awakenings was recorded, which was significant at the 7.5 g and 9.0 g/night Xyrem doses.
- Across doses, a non-significant decrease in Stage 1 sleep was observed, while the amount of Stage 2 sleep remained unchanged.
- An acute increase in REM sleep time was demonstrated with the initial 4.5 g treatment, with subsequent dose related significant decreases in total REM sleep time at all 4 doses.
- No dose related change in total sleep time was observed; however, a significant decrease in total sleep time was found following 4.5 g/night dosing for 4 weeks.
- The number of shifts in sleep stage demonstrated a decreasing trend at doses greater than 4.5 g/night but a significant decrease was recorded only following 7.5 g/night Xyrem.
- A significant dose-dependent increase in sleep latency was observed across all 4 doses.

9.11.2 Objective And Subjective Measures Of Daytime Sleepiness

- The administration of Xyrem produced a significant increase in daytime sleep latency as measured by the MWT. This dose-dependent increase averaged 3.7 minutes ($p = 0.038$) after 4 weeks of 4.5 g nightly that further increased to a mean improvement of 6.1 minutes ($p < 0.001$) following the nightly 9.0 g dose. This measured response is additive to that produced by concomitant stimulant dosing.
- The presence of sleep-onset REM periods during MWT, which occurred in 18 of 21 patients (86%) at baseline, decreased to 13 of 21 patients (62%)following 4 weeks of 4.5 g Xyrem nightly. Sleep-onset REM periods further decreased to 6 of 20 patients (30%) following the 9.0 g dose.

- The Epworth Sleepiness Scale total score significantly decreased in a dose-dependent manner by the nightly administration of Xyrem. The median total score of 20 at baseline improved by 2 points following the 4.5 g dose regimen ($p<0.001$), increasing across the dose range up to 7 points after the 9.0 g dose ($p<0.001$).

9.11.3 Overall Symptoms Of Narcolepsy

The patients in the current trial reported substantial improvements in subjectively determined (by NSA recording of) daytime narcolepsy symptoms including the incidence of cataplexy attacks, the number of inadvertent naps as well as decreased daytime sleepiness, an increased the ability to concentrate and a perception of overall improvement in their narcolepsy while taking nightly doses of Xyrem.

9.12 Reviewer's Comments

Since the study was uncontrolled and open-label in design, it does not lend itself to any conclusions regarding the efficacy of Xyrem® on polysomnogram variables, or on the other outcome measures used for this study. The results of this study cannot therefore be used to support any statements in labeling about the efficacy of Xyrem®.

10. Study OMC-SXB-21

The final report of this efficacy study was submitted on 12/16/00, i.e., after the original NDA submission.

10.1 Tabular Summary

Study #	OMC-SXB-21 Orphan Medical	
Design	Randomized, double-blind, placebo-controlled, parallel-arm, randomized withdrawal study after long-term open label treatment	
Duration	2 weeks (withdrawal phase)	
Study Arms	GHB	Placebo
Number receiving study drug	26	29
Number completed	26	29
Main inclusion criteria	Continuous treatment with GHB for narcolepsy for 6 to 35 months	
Primary outcome measures	Total number of cataplexy attacks	
Main efficacy analysis (statistically significant results)	GHB superior to placebo, based on ANCOVA ($p < 0.001$)	

10.2 Title

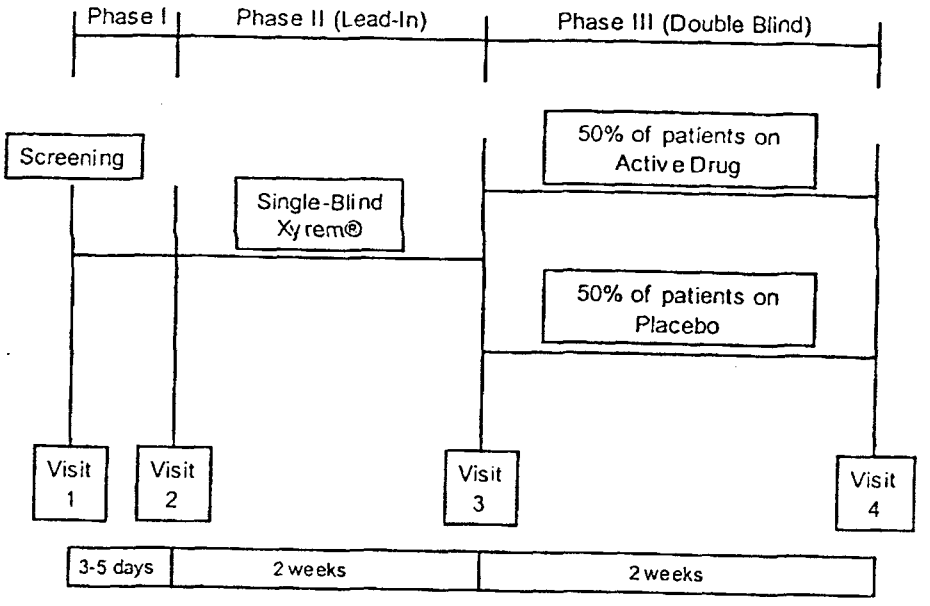
A Randomized, Double-Blind, Placebo-Controlled, Multi-Center Trial To Assess The Long-Term Efficacy Of Orally Administered Xyrem® (Sodium Oxybate) When Compared To Placebo

10.3 Objective

To provide evidence for the long-term efficacy of Xyrem® based upon the return of cataplexy symptoms upon cessation of a minimum of 6 months of open label treatment with active drug

10.4 Design

The design of the study is schematically summarized below



10.5 Duration

4 weeks (2 weeks of a double-blind withdrawal phase)

10.6 Sample Size

60 patients, with 30 in each treatment group in Phase 3, of the study will be included in the trial

10.7 Selection

10.7.1 Key Inclusion Criteria

- Informed consent
- Age ≥ 16 years
- Willing and able to complete the entire trial
- At least 5 cataplexy attacks per week prior to receiving any treatment (tricyclic antidepressants, selective serotonin uptake inhibitors, or Xyrem®) for cataplexy
- If female must be
  - Surgically sterile OR
  - 2 years post-menopausal OR
  - If of child bearing potential must be using a medically accepted means of birth control and must agree to continue such treatment for the duration of the study
- Treated continuously for the symptoms of narcolepsy with Xyrem® for at least 6 months, and not more than 3.5 years
- Willing to not operate a car or heavy machinery if the clinical investigator feels such a restriction is warranted
- Adequate support for the duration of the trial

10.7.2 Key Exclusion Criteria

- Unstable diseases in any body system, other than narcolepsy, which would place the patient at risk or compromise the trial objectives
- Psychiatric disorders that would preclude participation in, or completion of, the trial
- History of substance abuse, as defined by DSM-IV, currently or within the past year
- Serum creatinine > 2 mg/dl; AST or ALT > 2 x upper limit of normal; serum bilirubin > 1.5 x upper limit of normal; pre-trial electrocardiogram results demonstrating a clinically significant arrhythmia or 2^{nd} or 3^{rd} degree A-V block; history of myocardial infarction within the past 6 months
- Any untreated disorder other than narcolepsy that could be considered a primary cause of excessive daytime sleepiness, including sleep apnea syndrome (criteria specified)
- Occupation requiring variable shift or routine night shift work
- Use of tricyclic antidepressants, selective serotonin uptake inhibitors or medications for cataplexy other than Xyrem® in the 30 days prior to Visit 1 of the study
- Clinically significant history of head trauma; previous invasive cranial surgery; seizure disorder; use of anticonvulsant medication

10.7.3 Concomitant Medications

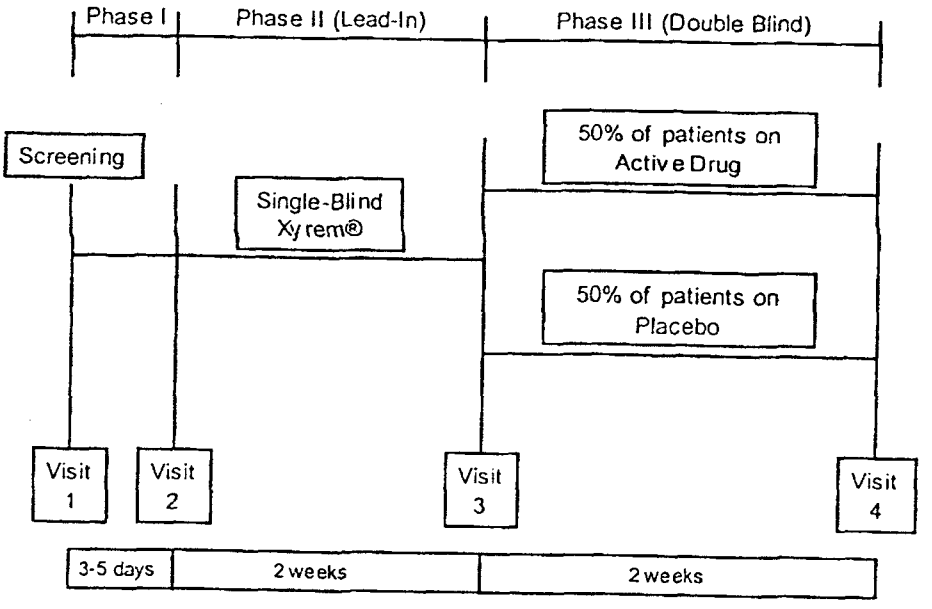
- The following medications are prohibited during the trial: selective serotonin uptake inhibitors and tricyclic antidepressants.
- Patients will be cautioned regarding the use of opioid analgesics and skeletal muscle relaxants
- Alcohol is prohibited during the trial
- Over-the-counter medications need careful review by the clinical investigator prior to use; non-sedating alternatives may be used wherever possible
- Stable doses of stimulant medication may be used to treat excessive daytime sleepiness as clinically indicated

10.8 Dosage

Previously established dose of Xyrem® ranging from 3 to 9 grams daily

10.9 Schedule

The study schematic is reproduced again here for convenience



- The visit schedule was as in the schematic above.
- The following were to be checked at Visit 1 alone: informed consent; selection criteria, medical history, cataplexy history prior to use of any medications, and "support systems".
- Physical examinations, including neurological examinations were to be performed at Visits 1 and 4
- Daily diaries were to be provided and/or checked at visits 2, 3, and 4. Diaries were to record cataplexy and adverse events.
- Concurrent medications, vital signs and adverse events were to be checked at every visit
- A pregnancy test were to be checked if applicable at Visit 1
- Routine hematology and chemistry were to be checked at Visits 1 and 4

10.10 Outcome Measures

10.10.1 Primary Efficacy

Frequency of cataplexy attacks

10.10.2 Secondary Efficacy

None stated

10.10.3 Safety

Adverse events, laboratory data

10.11 Safety Monitoring

Vital signs, safety laboratory tests, and physical examinations.